

0300

12



OIPÉ

RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/09/802,208B

TIME: 20:49:29

Input Set : N:\AMC\I802280b.raw

Output Set: N:\CRF4\09232002\I802208B.raw

ENTERED

```

1 <110> APPLICANT: Parrott, Wayne
2   LaFayette, Peter
3   Kane, Patrick
4 <120> TITLE OF INVENTION: Arabitol or Ribitol As Positive Selectable Markers
5 <130> FILE REFERENCE: UGA-855R
C--> 6 <140> CURRENT APPLICATION NUMBER: US/09/802,208B
C--> 7 <141> CURRENT FILING DATE: 2001-03-08
8 <160> NUMBER OF SEQ ID NOS: 5
9 <170> SOFTWARE: PatentIn version 3.0
11 <210> SEQ ID NO: 1
12 <211> LENGTH: 1848
13 <212> TYPE: DNA
14 <213> ORGANISM: Escherichia coli
15 <400> SEQUENCE: 1
16   atgaacgaac aatttacatg gctgcacatc gggttagggt cttttcatcg cgcacatcag      60
17   gcgtggatc tacaccgttt gcaggtagtg ggcgataaac gctggagcat tgctgcgggc      120
18   aatattcgta atgatgctga acatgtcgta caggcactca gtgcacagaa aggtcgctat      180
19   gtgctggaaa ccgtcagccc ggaaggggta agcgaatatg aagagatcac ctcaattcag      240
20   aagttgatac cgtggcaggc agatttaca ccgctgattg ctgaaggggc agatccgaag      300
21   acaaaagtga ttgctttcac cgtcaccgaa ggcgggtact acctgaatac cagtcacaaa      360
22   ctggaagtta acaatcctga tttagcggca gatcttaaag ggggatgcaa aacaatttac      420
23   ggtgttatta cccgtatcct cgaagcgcgt atggcaaata acgccggacc actaaccctg      480
24   atgaacgaac aatttacatg gctgcacatc gggttagggt cttttcatcg cgcacatcag      540
25   gcgtggatc tacaccgttt gcaggtagtg ggcgataaac gctggagcat tgctgcgggc      600
26   aatattcgta atgatgctga acatgtcgta caggcactca gtgcacagaa aggtcgctat      660
27   gtgctggaaa ccgtcagccc ggaaggggta agcgaatatg aagagatcac ctcaattcag      720
28   aagttgatac cgtggcaggc agatttaca ccgctgattg ctgaaggggc agatccgaag      780
29   acaaaagtga ttgctttcac cgtcaccgaa ggcgggtact acctgaatac cagtcacaaa      840
30   ctggaagtta acaatcctga tttagcggca gatcttaaag ggggatgcaa aacaatttac      900
31   ggtgttatta cccgtatcct cgaagcgcgt atggcaaata acgccggacc actaaccctg      960
32   ctgaattgcg ataacgtgcg ccataatggt gaacgtttcc atgatggcct ggttgagttt     1020
33   ctccagctaa ctggcaaaca ggatgtcatc gactggctga gtacaaatac cacttgcccg     1080
34   aataccatgg ttgaccgcat tacgcctcgt ccggcagcag aacttccggc acggatcaag     1140
35   gctcaaacgg gtattgccga taaagcgccg gtaatgggcg aaacctttat ccagtgggtc     1200
36   gtggaagata atttccgtga tgtccgtccg gcaactggaga aggtcggtgt cgaactgggtg     1260
37   gcgtcggtaa tcccctatga agaggcgaag attcgcattc ttaactcttc acacagttgc     1320
38   atcgctggg caggtacgtt aatcggtcaa aaatatatcc acgaaagcac aatgaccgat     1380
39   tttatctatc agattgccga ccgtacgtg acggaagatg tcattccttg cttgggcgat     1440
40   aacggtatcg atttgccaac ctaccgtgat gttgtactca agcgttttac caatccacat     1500
41   atccaggaca ccaaccaacg cgtcgcgtcg gatggtttct cgaaaattcc ggcatgatt     1560
42   gccccacac tgcgagagtg ctaccagcga ggcgttcgcc cgaatgccac gccatgtta     1620
43   cctgcactgt tttacgtatt catggagcag tggcatcacg gcaaaactgcc ctatgaatat     1680
44   caggatggca tccttgatgc accagctgtc catgcaatgt tacagtctgc cgatcccgtc     1740

```

RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/09/802,208B

TIME: 20:49:29

Input Set : N:\AMC\I802280b.raw

Output Set: N:\CRF4\09232002\I802208B.raw

```

45      gctgtttatg ccagtgataa agcgctgttt ggcgatttaa ccgaacgtga agattttgcc 1800
46      gcgttggtgc gcgaaaaaat cgctgacgtc tacgcgttaa ttaactaa 1848
48 <210> SEQ ID NO: 2
49 <211> LENGTH: 3971
50 <212> TYPE: DNA
51 <213> ORGANISM: Escherichia coli
52 <220> FEATURE:
53 <221> NAME/KEY: misc_feature
54 <222> LOCATION: (96)..(848)
55 <223> OTHER INFORMATION: ribitol dehydrogenase coding region
56 <220> FEATURE:
57 <221> NAME/KEY: misc_feature
58 <222> LOCATION: (859)..(2463)
59 <223> OTHER INFORMATION: ribitol kinase coding region
60 <220> FEATURE:
61 <221> NAME/KEY: misc_feature
62 <222> LOCATION: (2565)..(3839)
63 <223> OTHER INFORMATION: ribitol transporter coding region
64 <400> SEQUENCE: 2
65      atcgattgag cagtttgctt cacacggcag ctaaattccc gttcagtgcg tgcaagcaaa 60
66      cagagactat aaattcgccc tggtaaaagg attatatgat gaatcaactct gtgccctcta 120
67      tgaatactcc ccttaatggc aaagttgcag ctatcaactgg cgctgcgtca ggtattggcc 180
68      tgcaatgtgc aaaaacgctg ctcgatgcag gagcaaaggt agtattgatt gaccgggaag 240
69      gcgacaaact gcacaagatt gtcgctgagt taggcgaaaa cgcgtagcgc ctgcaactcg 300
70      atctcttcaa taatcagcaa gtcgataaca tgctggcgga cattatcgaa ctggcggtg 360
71      ggctggatat ttttcatgcc aatgcaggcg cttatatattg cggcccagtg gctgaagggtg 420
72      atccagatgt ctgggatcgt gtgttaaata tgaatataaa tgcggcggtt cgctgtgtcc 480
73      gtgcagtcct gccgcataat attgcgcaga ggtcgggcga tataattttt accagttcca 540
74      tcgcgggcgt cgttccggtt atctgggaac cgatctacac cgcgtcctaaa ttigccgttc 600
75      aggcattcgt acacactacc cgccggcagg tttctcaata tggcgtgcgt gtgggtgcgg 660
76      tgctgccagg accagtagtc actgccctgc ttgatgactg gccaaaagcc aaaatggaag 720
77      aagccctggc aaatggtagc ctgatgcaac cgattgaagt ggcggaatca gtattgttta 780
78      tggtgaccgc ctcgaaaaat gtcaccgtgc gagatttagt gatcctgcct ggcagtgtcg 840
79      atctgtaagg gcgcaatcat gacaataacc aaaaccgtta ttggtgttga tgtgggatca 900
80      ggcagtgtcc gcgcgggat ttttgatctc aacggatctc tgctatccca tgccacagaa 960
81      aaaatcacga ctacgcggcg cagcgggaagc cgcgtggaac agtccagcca ggagatctgg 1020
82      caggcgggtct gttcatgtat tcgtaatgcg ctactcttgg cagacgtttg tgcacaaagt 1080
83      gtggcaggca tcggttttga tgccacctgt tctctgggtg tactggataa aaacggtgat 1140
84      ccattgcctg tcagcccggg aggagatgca aagcaaaata tcattgtgtg gatggatcac 1200
85      cgcgccaccg aacaagcaga gcgaatcaat gccactcacc atccggtgct gaactacgtc 1260
86      ggtggtaaaa tttgcctga aatggaaaca ccgaaaattc tctggctgaa agaaaatatg 1320
87      ccagagatct acgaacgtgc cggacaattt ttgatcttgg ccgattttct gacctggcgg 1380
88      gctaccggtg atttagcgcg ttacgtatgc actgttacct gtaaattggac gtggctggca 1440
89      catgaaaatc gctgggatcc agattatttc cgcaccattg gccttgacga gtttagcggt 1500
90      gaagatttta ttgcattgg tcatcatatt gtttctcccg gaacaccttg tggaaatggt 1560
91      ttaacagcac aagccgcggc agagatggga ttactccccg gcacacctgt cgctgtaggg 1620
92      ttaattgatg ctacgcgtgg tggcatcggt acggttggcg tagaagggtg agcgtgaac 1680
93      aatctcgcgt atgttttcgg cacttcttca tgcaccatgg catctaccac ttctccctcg 1740
94      tttgtaccgg gtgtctgggg gccgtattac agtgcgatgg ttccaggtct gtggttagtt 1800

```

RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/09/802,208B

TIME: 20:49:29

Input Set : N:\AMC\I802280b.raw

Output Set: N:\CRF4\09232002\I802208B.raw

```

95      gaagggcgggc aaagtgtctgc gggagcagct attgaccagc tacttgattt ccatccggct 1860
96      gttgaagaag ctgcgcgaaat ggcacaacgt gtgaatcagc ccctccccgt ctggcttgct 1920
97      gatcgaatcc tcgaaaaaac ggcgcaacca tcagatgtctg tcgccctggc gaaagggcta 1980
98      cacgtgggtgc cggaatttct cggaatcgc gcccccttcg cagatcctca tgccagagcg 2040
99      gtaatttggtg gcctgggtat ggagcgagat ctggataatt tactcgccct gtatatcgct 2100
100     ggattatgcg gaattggtta tggctctgcgc caaattctcg acgctcaaac agcgcagggg 2160
101     gtagtgagta aaaatatcgt tattagcggc ggtgcccggc agcatccact ggtacgacaa 2220
102     attctggcag atacctgcgg tattccggtc attaccacgc aatgctgcga accggtttta 2280
103     ttaggtctcg ccattcttgg tgctgtcgcc ggaaatattg caccttctgt tggcgaagcg 2340
104     atgcaacaat tcacccatgt ggataaatat tattatccgc aagaacgcta tcaatctctt 2400
105     catcatcgtc gatatgaggc ctataagcag ttgcagcata ctgcaaaatt actcagagac 2460
106     taattaacca gccgcctgac gctgttttca ggcaatcact aataacgact cactccggta 2520
107     atatcccgga gtgcattcat ctgcacccta aaaacgaggt ttatatgtcc agaaaataata 2580
108     aacagtgggt gggtttgcca ctgcatctga tatggggata tatcgccatc gcagtattta 2640
109     tgactgggtg tggtttcgaa ctgcgattct tatcgcata tattaaatcg ttaggcttca 2700
110     cacctgcgga agcctctttt gcctttacgc tctacggcct ggcggtcgcc ctttccgcct 2760
111     gggtttctgg ggtagtagcg gaaatcatca cgccgcaaaa aaccatgctg attggttttg 2820
112     tcctatgggt cgtattccat gttctgtttc tggcttttgg attaggacag gcaaactatg 2880
113     ggttaatcct gctgttttac gggattcgtg gtttagcgta tccgctatct ctctactcat 2940
114     ttatcgctgt tattattcat aacgtgcgta gcgaaaattc cagttctgct ctgggttggt 3000
115     actgggcggg atattcagta gggatcgggt ttgctggcag ttatatccc agttttacga 3060
116     taccgattat gggtgaaatg ggaaccttat ggctggcact ggctgtctgc tttgctggcg 3120
117     gtgtcattgc catgatctcc ttgcgtcatg ttaaaacgcc tggacatatg cataatttaa 3180
118     ctccccgtga gaagtttgca gaattaagtc gggcagtaac tttactttat accaaccgca 3240
119     atatttttct ctccagtatt gtgcgcatta taaataacct atcgttattc ggttttgcg 3300
120     tcattatgcc aatgatgttt gtggatgaac tgggattcac cacctctgaa tggttgcagg 3360
121     tctgggcggc atttttcttt accactatct tctctaata tttttggggg attgtggcag 3420
122     aaaaaatggg ctggatgcgt gttattcgct ggtttgggtg cctcgggatg gcagcatcaa 3480
123     gtttagcggt ttactacatg ccgcaatact ttggtcacaa ctactggatg gcaatgattc 3540
124     cggcgattgc tctgggaaat tttgttgctg catttggtgc gatggccgct gtcttcccg 3600
125     cactggaacc aaaacacaaa ggtgctgcaa tctcggttta caacctctct gcgggtatgt 3660
126     ctaacttcct ggtccggcga attgccgtgg tgttattacc gtggttttagc actatcggtg 3720
127     tggtcattgc ctatacagca ttgtatctat tggcctttgt cctttgcgca ttcattcgcg 3780
128     ttgagcagcc aggattcagt tctgcgccag tgactgagaa ggcattgaat atctcctgaa 3840
129     aaacgaaacg catcaggcac tcacccctct cctcatggga gaggatgatt tcacatcagg 3900
130     caatagtgac tttgttatcg agataaacgt cctgcacggc gttaatcagt ttcacgcccgt 3960
131     cagccatcga t 3971

```

133 <210> SEQ ID NO: 3

134 <211> LENGTH: 250

135 <212> TYPE: PRT

136 <213> ORGANISM: Escherichia coli

137 <400> SEQUENCE: 3

```

138      Met Met Asn His Ser Val Pro Ser Met Asn Thr Pro Leu Asn Gly Lys
139      1              5              10              15
140      Val Ala Ala Ile Thr Gly Ala Ala Ser Gly Ile Gly Leu Gln Cys Ala
141      20              25              30
142      Lys Thr Leu Leu Asp Ala Gly Ala Lys Val Val Leu Ile Asp Arg Glu
143      35              40              45
144      Gly Asp Lys Leu His Lys Ile Val Ala Glu Leu Gly Glu Asn Ala Tyr

```

RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/09/802,208B

TIME: 20:49:29

Input Set : N:\AMC\I802280b.raw

Output Set: N:\CRF4\09232002\I802208B.raw

```

145          50          55          60
146 Ala Leu Gln Leu Asp Leu Phe Asn Asn Gln Gln Val Asp Asn Met Leu
147 65          70          75          80
148 Ala Asp Ile Ile Glu Leu Ala Gly Gly Leu Asp Ile Phe His Ala Asn
149          85          90          95
150 Ala Gly Ala Tyr Ile Gly Gly Pro Val Ala Glu Gly Asp Pro Asp Val
151          100          105          110
152 Trp Asp Arg Val Leu Asn Leu Asn Ile Asn Ala Ala Phe Arg Cys Val
153          115          120          125
154 Arg Ala Val Leu Pro His Met Ile Ala Gln Arg Ser Gly Asp Ile Ile
155          130          135          140
156 Phe Thr Ser Ser Ile Ala Gly Val Val Pro Val Ile Trp Glu Pro Ile
157 145          150          155          160
158 Tyr Thr Ala Ser Lys Phe Ala Val Gln Ala Phe Val His Thr Thr Arg
159          165          170          175
160 Arg Gln Val Ser Gln Tyr Gly Val Arg Val Gly Ala Val Leu Pro Gly
161          180          185          190
162 Pro Val Val Thr Ala Leu Leu Asp Asp Trp Pro Lys Ala Lys Met Glu
163          195          200          205
164 Glu Ala Leu Ala Asn Gly Ser Leu Met Gln Pro Ile Glu Val Ala Glu
165          210          215          220
166 Ser Val Leu Phe Met Val Thr Arg Ser Lys Asn Val Thr Val Arg Asp
167 225          230          235          240
168 Leu Val Ile Leu Pro Gly Ser Val Asp Leu
169          245          250
171 <210> SEQ ID NO: 4
172 <211> LENGTH: 534
173 <212> TYPE: PRT
174 <213> ORGANISM: Escherichia coli
175 <400> SEQUENCE: 4
176 Met Thr Ile Thr Lys Thr Val Ile Gly Val Asp Val Gly Ser Gly Ser
177 1          5          10          15
178 Val Arg Ala Gly Ile Phe Asp Leu Asn Gly Ser Leu Leu Ser His Ala
179          20          25          30
180 Thr Glu Lys Ile Thr Thr Thr Arg Arg Ser Gly Ser Arg Val Glu Gln
181          35          40          45
182 Ser Ser Gln Glu Ile Trp Gln Ala Val Cys Ser Cys Ile Arg Asn Ala
183 50          55          60
184 Leu Thr Leu Ala Asp Val Cys Ala Gln Ser Val Ala Gly Ile Gly Phe
185 65          70          75          80
186 Asp Ala Thr Cys Ser Leu Val Val Leu Asp Lys Asn Gly Asp Pro Leu
187          85          90          95
188 Pro Val Ser Pro Glu Gly Asp Ala Lys Gln Asn Ile Ile Val Trp Met
189          100          105          110
190 Asp His Arg Ala Thr Glu Gln Ala Glu Arg Ile Asn Ala Thr His His
191          115          120          125
192 Pro Val Leu Asn Tyr Val Gly Gly Lys Ile Ser Pro Glu Met Glu Thr
193 130          135          140
194 Pro Lys Ile Leu Trp Leu Lys Glu Asn Met Pro Glu Ile Tyr Glu Arg

```

RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/09/802,208B

TIME: 20:49:29

Input Set : N:\AMC\I802280b.raw

Output Set: N:\CRF4\09232002\I802208B.raw

195	145	150	155	160
196	Ala Gly Gln Phe Phe Asp Leu Ala Asp Phe Leu Thr Trp Arg Ala Thr			
197		165	170	175
198	Gly Asp Leu Ala Arg Ser Val Cys Thr Val Thr Cys Lys Trp Thr Trp			
199		180	185	190
200	Leu Ala His Glu Asn Arg Trp Asp Pro Asp Tyr Phe Arg Thr Ile Gly			
201		195	200	205
202	Leu Ala Glu Leu Ala Asp Glu Asp Phe Ile Arg Ile Gly His His Ile			
203		210	215	220
204	Val Ser Pro Gly Thr Pro Cys Gly Asn Gly Leu Thr Ala Gln Ala Ala			
205		225	230	235
206	Ala Glu Met Gly Leu Leu Pro Gly Thr Pro Val Ala Val Gly Leu Ile			
207		245	250	255
208	Asp Ala His Ala Gly Gly Ile Gly Thr Val Gly Val Glu Gly Gly Ala			
209		260	265	270
210	Leu Asn Asn Leu Ala Tyr Val Phe Gly Thr Ser Ser Cys Thr Met Ala			
211		275	280	285
212	Ser Thr Thr Ser Pro Ser Phe Val Pro Gly Val Trp Gly Pro Tyr Tyr			
213		290	295	300
214	Ser Ala Met Val Pro Gly Leu Trp Leu Val Glu Gly Gly Gln Ser Ala			
215		305	310	315
216	Ala Gly Ala Ala Ile Asp Gln Leu Leu Asp Phe His Pro Ala Val Glu			
217		325	330	335
218	Glu Ala Arg Glu Met Ala Gln Arg Val Asn Gln Pro Leu Pro Val Trp			
219		340	345	350
220	Leu Ala Asp Arg Ile Leu Glu Lys Thr Ala Gln Pro Ser Asp Ala Val			
221		355	360	365
222	Ala Leu Ala Lys Gly Leu His Val Val Pro Glu Phe Leu Gly Asn Arg			
223		370	375	380
224	Ala Pro Phe Ala Asp Pro His Ala Arg Ala Val Ile Cys Gly Leu Gly			
225		385	390	395
226	Met Glu Arg Asp Leu Asp Asn Leu Leu Ala Leu Tyr Ile Ala Gly Leu			
227		405	410	415
228	Cys Gly Ile Gly Tyr Gly Leu Arg Gln Ile Leu Asp Ala Gln Thr Ala			
229		420	425	430
230	Gln Gly Val Val Ser Lys Asn Ile Val Ile Ser Gly Gly Ala Gly Gln			
231		435	440	445
232	His Pro Leu Val Arg Gln Ile Leu Ala Asp Thr Cys Gly Ile Pro Val			
233		450	455	460
234	Ile Thr Thr Gln Cys Cys Glu Pro Val Leu Leu Gly Ser Ala Ile Leu			
235		465	470	475
236	Gly Ala Val Ala Gly Asn Ile Ala Pro Ser Val Gly Glu Ala Met Gln			
237		485	490	495
238	Gln Phe Thr His Val Asp Lys Tyr Tyr Pro Gln Glu Arg Tyr Gln			
239		500	505	510
240	Ser Leu His His Arg Arg Tyr Glu Ala Tyr Lys Gln Leu Gln His Thr			
241		515	520	525
242	Ala Lys Leu Leu Arg Asp			
243		530		

VERIFICATION SUMMARY

DATE: 09/23/2002

PATENT APPLICATION: US/09/802,208B

TIME: 20:49:30

Input Set : N:\AMC\I802280b.raw

Output Set: N:\CRF4\09232002\I802208B.raw

L:6 M:270 C: Current Application Number differs, Replaced Current Application Number

L:7 M:271 C: Current Filing Date differs, Replaced Current Filing Date